AD 697 422

STOCHASTIC PROGRAMS WITH RECOURSE: A SURVEY 1

Roger J.-B. Wets

Boeing Scientific Research Laboratories
Seattle, Washington

August 1969

BOEING SCIENTIFIC RESEARCH LABORATORIES DOCUMENT D1-82-0982

BOEING
SCIENTIFIC RESEARCH LABORATORIES

AD 697422

Stochastic Programs with Recourse: A Survey I.

Roger J.-B. Wets

D1-82-0882

STOCHASTIC PROGRAMS WITH RECOURSE: A SURVEY I.

by

Roger J.-B. Wets

Mathematical Note No. 614

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

August 1969

Abstract

~~We~~ study the properties of the deterministic equivalent program of a stochastic program with recourse. After a brief discussion of the place of the stochastic programming model in the realm of stochastic optimization and the definition of the problem under consideration, ~~the three following sections are devoted~~ to the characterization of feasible solutions. ~~The two last sections~~ examine the various properties of the objective functions (dependent on or independent from the type of distribution of the random elements).

1. __Introduction__. In the first part of this survey, we shall be particularly concerned with the properties of the deterministic equivalent program of a stochastic program with recourse. We shall restrict our attention to the two-stage linear problem with *fixed* recourse. These limitations have for principal purpose to simplify substantially the presentation. However, whenever it will seem to be appropriate we shall indicate the extent to which certain of these results can be generalized.[+] On the other hand, although the title indicates that this paper is essentially a survey, a fair portion of the results mentioned here have never appeared in print; most of these can be considered as sharpened versions of propositions which have appeared previously or as a different approach to known results, which allows for a somewhat more satisfactory treatment. A few results are genuinely original.

We do not intend to give an historical account of the various developments in this field; however, we shall try to sketch the evolution one can observe in the statement of the problem as well as the various techniques which have been used to resolve some of these difficulties. The following parts of the survey will be devoted to some characterizations of the optimal solutions of stochastic programs, to the nonlinear and multiple recourse formulations and to some computational experiences as well as a few selected applications.

[+]Some generalizations are so trivial in nature that they might remain unmentioned.

Before we give a precise mathematical definition of the problem, it might be worthwhile to place this class of problems in a somewhat larger context. Consider a decision maker faced with various alternatives. His decisions are subject to a certain number of constraints and his goal is to optimize a given criterion. However, at the time of his selection of one of the alternatives, some of the parameters of the problem are unknown to him. The amount of information which is given about these parameters divides this family of decision making problems into two main classes to which the economists refer as: decision making under uncertainty and decision making under risk. In this first class of problems, essentially nothing (or very little) of a quantitative nature is available to the decision maker. In the second class, it is assumed that the decision maker is given a description of these unknown parameters in terms of a well determined probability law. (Various papers have studied models where some partial knowledge of this probability law is available. In [Miyasawa, 1968], and [Avriel and Williams, 1969] one can find a discussion of information structures in stochastic programming problems.)

The term *Stochastic Programming* will be reserved for those programming problems which fit in this second class. In the literature one finds various papers dealing with mathematical programs whose parameters, or some of their parameters, are random, but which by our definition may have little to do with stochastic programming. Some of these papers discuss various problems the importance of which ranges from amazing mathematical puzzles up to some important and sometimes difficult questions whose solution would most definitely enhance our chances of solving "difficult" stochastic programs. For example, a complete solution to the distribution problem:

> Find the distribution function of $Q(A,b,c)$
> where $Q(A,b,c) = \{\text{Min } cx \mid Ax = b, x \geq 0\}$
> and A,b,c are random matrices,

is neither expected, nor is it necessary for solving a large class of stochastic programs; however, it is obvious that any additional insight one might gain in this question could prove very worthwhile.

Rather than giving an axiomatic definition of a model for stochastic programs, as can be done for linear programs [Dantzig, 1963, Ch. 3], we shall limit ourselves to some of the characteristics of stochastic optimization problems. In some sense we shall try to point to those properties which constitute the essence of stochastic optimization models. First, perhaps, the *sequential nature* of the problem, i.e., the decision, is selected before the random elements, say ξ, of the problem can be observed. Second, the solution is the selection of *one particular decision*, say x in $\mathcal{X}$, where $\mathcal{X}$ denotes the set of alternatives. Third, the *actual value* $z(x,\xi)$ of a given decision can only be determined after observing ξ. Fourth, it is assumed that the decision maker has a *definite attitude towards risk* described by some function, say u, which usually has all the properties of the standard utility function in mathematical economics. In particular, this implies that the criterion for optimization can always be expressed in terms of the maximization (or minimization) of the expectation of utility. Finally, it is assumed that the probability law of ξ is completely determined once a particular x is selected, i.e., we are *given the probability space* $(\Xi_x, \mathcal{G}_x, \mu_x)$ on which ξ is defined.

Thus a very abstract formulation of a stochastic optimization problem could be:

$$\text{Given } [(\Xi_x, \mathcal{G}_x, \mu_x), \mathcal{X}, z, u],$$
$$\text{find } \inf E\{u(z(x,\xi))\}$$

We do not intend to develop any further an abstract theory of optimization, nor do we intend to make any specific use of the above model. Our purpose in formulating a fairly abstract stochastic optimization problem is that it might serve as a point of reference whenever we study any particular model for stochastic optimization. Many problems studied in mathematical economics, operations research, system engineering and control theory, fit into this class of optimization problems: e.g., Markov Programming, Inventory Problems, Stochastic Optimal Control Problems,... Determining to which category a given stochastic optimization problem belongs is irrelevant. In practice, these categories have considerable overlap, e.g., the newsboy problem[†] which is probably the first example of a stochastic programming problem found in the literature, is usually, and rightfully so, thought of as belonging to that class of problems dealt

---

[†]Briefly: Given a distribution of demand, a purchase cost, a selling price and given that no unsold papers can be returned, a newsboy seeks the optimal number of papers to order so as to maximize his expected profit.

with in inventory theory.

The first formulation of stochastic programs with recourse was given by G. Dantzig [Dantzig, 1955] who called them linear programs under uncertainty, and by E. Beale [Beale, 1955] who used the name of linear programs with random coefficients. Beale's formulation can be viewed as an n-dimensional generalization of the newsboy problem. Dantzig's formulation is in some sense somewhat more general, but still includes many of the particular properties of the n-dimensional newsboy problem. (See in particular the application [Ferguson and Dantzig, 1956] which motivated Dantzig's work.) As already mentioned above the actual formulation of the problem has not remained static; this was essentially due to the necessity of widening the range of possible applications as well as to the desire for somewhat more rigorous mathematical foundations for the theory. Although we will not pursue this objective here, this "grown-up" model for stochastic programming also allows for a more interesting economic analysis.

2. Statement of the Problem. Let $\xi = (c(\xi), q(\xi), p(\xi), T(\xi))$ be a random vector defined on a probability space $(\Xi, \mathcal{F}, \mu)$ where Ξ is a Borel subset of R^N, $N = (n+1)(\bar{m}+1) + \bar{n} - 1$, $\mathcal{F}$ is a σ-algebra containing the Borel subsets of Ξ, μ is a probability measure defined on Ξ. The coordinates of a point ξ in Ξ is a collection of four matrices c, q, p and T of dimension $1 \times n$, $1 \times \bar{n}$, $\bar{m} \times 1$ and $\bar{m} \times n$ respectively, i.e., c, q, p, and T are projections of ξ. By F we denote the distribution function associated with μ and by $\tilde{\Xi}$ the support of

the distribution of ξ, i.e., the smallest closed subset of R^N of measure 1.

Definition (2.1). If for all i,j,k the random functions $c_j(\xi)$, $q_j(\xi)p_i(\xi)$ and $q_j(\xi)t_{ik}(\xi)$ have first moments, then the random variable ξ is said to satisfy a *(weak) covariance condition.* We shall assume that such a condition is satisfied by the random elements of the problem.

From a practical viewpoint this condition might seem difficult to verify, however it is easy to see that if ξ has variance, then assumption (2.1) is automatically satisfied. Similarly, if either q or p and T are fixed, then it suffices for ξ to have first moments to satisfy the covariance condition (2.1). Later when we shall refer to stochastic programs satisfying a weak covariance condition, we shall always assume that its random elements satisfy (2.1) or some other condition which is weaker than (2.1), but would be sufficient to make the statement valid.

A *stochastic programming problem with fixed recourse* can be formulated as:

$$(2.2) \quad \text{Find} \inf_{\substack{x\geq 0 \\ Ax=b}} z(x) = E_\xi\{c(\xi)x + \min_{y\geq 0}[q(\xi)y \mid T(\xi)x + Wy = p(\xi)]\}$$

where A, W, and b are fixed matrices of size $m \times n$, $\bar{m} \times \bar{n}$, and $m \times 1$ respectively, and E_ξ denotes expectation with respect to ξ. We speak here of fixed recourse by opposition to the case when the matrix W is also random. This more general case has been discussed in [Walkup and Wets, 1967-b].

Without loss of generality, we can assume that W has full rank, since otherwise there is a linear non-singular transformation of the system of equations $T(\xi)x + Wy = p(\xi)$ which would render some of the rows of W identically zero; in which case the corresponding constraints either generate some deterministic restrictions on the decision variable x which could have been included in the constraints $Ax = b$, or they constitute a genuine system of stochastic equations whose solution depends non-trivially on ξ, i.e., no particular vector x can be found which would satisfy these equations with probability 1. In this last eventuality, the problem would be infeasible.

As a function of x and ξ, which we shall denote by $Q(x,\xi)$, the problem $\{\operatorname*{Min}_{y\geq 0} q(\xi)y \mid Wy = p(\xi) - T(\xi)x\}$ might be infeasible or unbounded below, in which cases we shall adopt the standard conventions and define $Q(x,\xi)$, i.e., the optimal value of the program to be $+\infty$ and $-\infty$. Thus the integrand of (2.1) might vary from $+\infty$ to $-\infty$ inclusively. Accordingly, the integral $E_\xi\{\cdot\} = \int \cdot \, d\mu$ is defined to be $+\infty$, if the integrand takes on the value $+\infty$ on a set of positive measure. It is $-\infty$ if the integrand is less than $+\infty$ almost everywhere and takes on the value $-\infty$ on a set of positive measure. If the integrand is finite almost everywhere, it corresponds to the Lebesgues-Stieltjes integral. This integral possesses essentially the same properties as the standard Lebesgues-Stieltjes integral except that subadditivity replaces the usual additivity property.

Due to the multistage nature of (2.2) it has been traditional and also convenient to write (2.2) as follows:

Without loss of generality, we can assume that W has full rank, since otherwise there is a linear non-singular transformation of the system of equations $T(\xi)x + Wy = p(\xi)$ which would render some of the rows of W identically zero; in which case the corresponding constraints either generate some deterministic restrictions on the decision variable x which could have been included in the constraints $Ax = b$, or they constitute a genuine system of stochastic equations whose solution depends non-trivially on ξ, i.e., no particular vector x can be found which would satisfy these equations with probability 1. In this last eventuality, the problem would be infeasible.

As a function of x and ξ, which we shall denote by $Q(x,\xi)$, the problem $\{\operatorname*{Min}_{y\geq 0} q(\xi)y \mid Wy = p(\xi) - T(\xi)x\}$ might be infeasible or unbounded below, in which cases we shall adopt the standard conventions and define $Q(x,\xi)$, i.e., the optimal value of the program to be $+\infty$ and $-\infty$. Thus the integrand of (2.1) might vary from $+\infty$ to $-\infty$ inclusively. Accordingly, the integral $E_{\xi}\{\cdot\} = \int \cdot \, d\mu$ is defined to be $+\infty$, if the integrand takes on the value $+\infty$ on a set of positive measure. It is $-\infty$ if the integrand is less than $+\infty$ almost everywhere and takes on the value $-\infty$ on a set of positive measure. If the integrand is finite almost everywhere, it corresponds to the Lebesgues-Stieltjes integral. This integral possesses essentially the same properties as the standard Lebesgues-Stieltjes integral except that subadditivity replaces the usual additivity property.

Due to the multistage nature of (2.2) it has been traditional and also convenient to write (2.2) as follows:

$$(2.3)\qquad \begin{array}{llll} \text{Find } \inf z(x) = E_\xi\{c(\xi)x + \text{Min } q(\xi)y\} & & & \\ \text{subject to} \quad Ax & & = b & \\ \qquad T(\xi)x + & Wy & = p(\xi) & \text{a.e.} \\ \qquad x \geq 0, & y \geq 0 & & \end{array}$$

where it is understood that the stochastic constraints $T(\xi)x + Wy = p(\xi)$ must be satisfied with probability 1 (a.e. denotes almost everywhere.) As we shall see, this last condition follows immediately from the definition of $E_\xi\{\cdot\}$. From a theoretical viewpoint it is sometimes convenient to amalgamate the fixed constraints $Ax = b$, $x \geq 0$ and the stochastic constraints; however, for practical purposes such as formulation and computational procedures it is clearly advantageous to separate them. We shall follow this practice here.

Problem (2.3) can be seen to have all the attributes of the class of stochastic optimization problems described in the introduction. The sequential nature of the problem [Dantzig, 1955, Introduction] is clearly illustrated: The decision process involves a choice of a decision x, then ξ occurs and is observed; finally a recourse action y is selected so as to satisfy the stochastic constraints. Since the decision y is completely determined by the selection of a given x and the occurrence of some ξ, the only *real* decision is the choice of x. The actual value of the decision is determined as soon as x and ξ are known, even though determining this value involves solving a linear program; the structure of the problem determines this value uniquely. The attitude towards risk of the decision maker is reflected by the linearity of the objective, i.e., it indicates neither risk aversion nor risk seeking. Finally, the definition itself of the problem guarantees that one has all the necessary information concerning the random elements of the problem.

3. The Deterministic Equivalent Program. As mentioned in the introduction, the purpose of the first part of this survey is to characterize the so-called deterministic equivalent program of problem (2.2). Let

$$z(x,\xi) = c(\xi)x + Q(x,\xi)$$

then

$$z(x) = E_{\xi}\{z(x,\xi)\} = E_{\xi}\{c(\xi)x + Q(x,\xi)\} = \bar{c}x + Q(x) \tag{3.1}$$

where $\bar{c}$ the expectation of $c(\xi)$ is finite, since assumption (2.1) implies that all the components of ξ are integrable. Also, in view of the definition of the function $Q(x,\xi)$ and the integral, the functions $z(x)$ and $Q(x)$ are well defined for all x in R^n as functions with range in the extended reals, provided $Q(x,\xi)$ is measurable. This is fairly easy to prove, see e.g., [Kall, 1967, Section 1, Satz 1] or [Walkup and Wets, 1967-b, Lemma (2.3)]. Note that among other things, it is possible for $Q(x)$ and thus $z(x)$ to be identically $+\infty$ or take on only the values $+\infty$ and $-\infty$. The problem

$$\begin{aligned} \text{Find } \inf z(x) &= \bar{c}x + Q(x) \\ Ax &= b \\ x &\geq 0 \end{aligned} \tag{3.2}$$

is known as the *deterministic equivalent program* of a stochastic program with recourse. Later on we shall give a more detailed form of the deterministic equivalent program, but before we do so, we need to probe somewhat deeper into the particularities of the function $Q(x)$. The next section is essentially devoted to the description of the region of finiteness of $Q(x)$. The properties that $Q(x)$ possess on this set receive further elaboration in Section 7.

4. Feasibility Region. The feasibility region of a mathematical program is usually defined to be the set of points satisfying the constraints. One can also define the feasibility region as the set of those points at which the objective is less than $+\infty$ accepting implicitly the convention that the objective is $+\infty$ outside the set determined by the constraints. In most practical situations these two definitions are equivalent. That this is also the case, for the class of problems under consideration, does not follow--as is usual--from the definition of the problem. In fact, the notion of feasibility is a very touchy problem, especially because the more natural definitions of feasibility do not seem to yield any grip on the computational aspects of the problem, whereas a more constructive type of definition seems to be a little too restrictive.

In the first papers dealing with stochastic programming this problem did not arise, because the models considered were such that some assumption, stated explicitly or present implicitly, eliminated this question. Let

$$K_1 = \{x \mid Ax = b, \; x \geq 0\}$$

be the set determined by the fixed constraints of the problem. If we assume *relatively complete recourse*, i.e., for all x in K_1 there always exists some $y \geq 0$ such that $Wy = p(\xi) - T(\xi)x$ for every ξ in R^N, then obviously the feasibility region is determined by K_1 alone. One can verify that the models considered by [Beale, 1955], [Dantzig, 1955], [Dantzig and Madansky, 1961], [Williams, 1963], [Williams, 1965], [Wets, 1966-a], and [El-Agizy, 1967], to name but a few, all implied relatively complete recourse. This assumption limits somewhat arbitrarily the range of possible applications, especially with respect to production models; and even though in most cases its verification might be easy, in general this might not be the case. During the last five or six years a certain number of papers [Wets, 1966-b], [Kall, 1967], [Dempster, 1968], and [Parrikh, 1967], to again mention only a few, started with approaching the problem of (i) establishing criteria for determining when the assumption of relatively complete recourse is satisfied, as well as (ii) considering a more general model for stochastic programming where this undesirable restriction has been removed. In Section 6, we shall discuss the problem of verifying the assumption of relatively complete recourse, as well as some related questions. For the time being we restrict ourselves to the problem we have defined and characterize as completely as possible its region of feasibility.

As indicated above, we can define the feasibility set as the intersection of the set determined by the fixed constraints K_1 and a set K_2 determined by the *induced constraints* generated by the stochastic constraints. We consider the following three candidates for K_2:

(i) $K_2^{\mu} = \{x \mid \text{with probability } 1 \ \exists\, y \geq 0 \text{ such that } Wy = p(\xi) - T(\xi)x\}$

$= \{x \mid Q(x,\xi) = +\infty \text{ with probability zero}\}$

That these two definitions of K_2^{μ} are equivalent follows immediately from the $+\infty$ and $-\infty$ conventions we have adopted for the definition of the function $Q(x,\xi)$.

(ii) $K_2^{s} = \{x \mid Q(x) < +\infty\}$

This set is known as the *strong feasibility set* since it clearly defines the region of finiteness of $Q(x)$ (unless $Q(x)$ takes on the value $-\infty$). These two first definitions of feasibility seem to lead to natural interpretations from a mathematical viewpoint as well as from an economic standpoint. The third definition

(iii) $K_2^{p} = \{x \mid \forall \xi \in \tilde{\Xi},\ \exists y \geq 0 \text{ such that } Wy = p(\xi) - T(\xi)x\}$

can be viewed as requiring that the stochastic constraints be satisfied for all points of the random elements which are "possible" where possible is defined as those points lying in the support of the distribution of ξ. As we shall see, this last definition allows for a more constructive approach to the description of the feasibility region, but from a purely mathematical viewpoint this "possibility" interpretation of the stochastic constraints is not as gratifying as the probabilistic interpretation associated with the first two definitions.

Although the following is not the case in general [Walkup and Wets, 1967-b], for the problem at hand, we have that:

Theorem (4.1). For a stochastic program with fixed recourse where ξ satisfies a weak covariance condition, such as condition (2.1), the relation

$$K_2^{\mu} = K_2^{P} = K_2^{s} \tag{4.2}$$

holds.

Proof. Since $\tilde{\Xi}$ is some set of probability 1, it follows that $K_2^P \subset K_2^{\mu}$. Thus to prove the first part of the theorem it suffices to show inclusion in the other direction. The following notation has proved to be very useful for problems of this type. Let

$$\text{pos } W = \{t \mid t = Wy,\ y \geq 0\} \tag{4.3}$$

be the positive hull spanned by the columns of W, i.e., the closed convex cone with apex at the origin generated positively by the points of $R^{\bar{m}}$ corresponding to the columns of W. Thus, we can write

$$K_2^{\mu} = \{x \mid p(\xi) - T(\xi)x \in \text{pos } W \text{ with probability } 1\}$$

and

$$K_2^{P} = \{x \mid p(\xi) - T(\xi)x \in \text{pos } W \text{ for all } \xi \text{ in } \tilde{\Xi}\}$$

If $x \in K_2^{\mu}$ then pos W contains a set, say Λ_x, of measure 1 with respect to the measure induced by the random vector $\ell_x(\xi) = p(\xi) - T(\xi)x$. Since pos W is closed, the closure of Λ_x is also contained in pos W and at fortiori so is $\tilde{\Lambda}_x$ the suppport of $\ell_x(\cdot)$. Since ℓ_x is a linear function of ξ (by definition $p(\cdot)$ and $T(\cdot)$ are canonical projections of ξ) it follows that $\tilde{\Lambda}_x$ and the closure of $\ell_x(\tilde{\ })$ coincide, i.e., $x \in K_2^P$. Thus $K_2^{\mu} = K_2^P$.

To show the remaining equality we first observe that the inclusion $K_2^{\mu} \supset K_2^{s}$ holds trivially from the definition of the integral. Since $K_2^{\mu} = K_2^{p}$ and $\tilde{\Xi}$ has measure 1, there remains only to be shown that for every x in K_2^{p}

$$Q(x) = \int_{\tilde{\Xi}} Q(x,\xi)\,d\mu < +\infty.$$

Let $\{W_{(i)}\}$ be a subcollection of the square non-singular submatrices of W of rank $\bar{m}$, such that the $\{W_{(i)}\}$ are distinct and such that $\{\text{pos } W_{(i)}\}$ constitutes a covering of pos W. Let $q(\xi)_{(i)}$ denote the subvector of $q(\xi)$ corresponding to the columns of W determining $W_{(i)}$. For every x in K_2^{p} and every ξ in Ξ, $p(\xi) - T(\xi)x$ belongs to some pos $W_{(i)}$ and since $W_{(i)}$ determines a feasible but not necessarily optimal basic solution for the recourse problem whose right-hand side is $p(\xi) - T(\xi)x$, we have that

$$Q(x,\xi) \leq q(\xi)_{(i)} W_{(i)}^{-1} [p(\xi)-T(\xi)x] = \nu_{(i)}(x,\xi).$$

The integral is isotone, thus

$$\int_{\Xi_{(i)}(x)} Q(x,\xi)\,d\mu \leq \int_{\Xi_{(i)}(x)} \nu_{(i)}(x,\xi)\,d\mu$$

where $\Xi_{(i)}(x) = \{\xi \,|\, p(\xi) - T(\xi)x \in \text{pos } W_{(i)}\}$. The right-hand side of this inequality is finite, since ξ satisfies the covariance condition (2.1). Let $\{\Sigma_{(j)}(x)\}$ be a finite partition of $\tilde{\Xi}$, where each $\Sigma_{(j)}(x)$ is obtained from the finite collection $\{\Xi_{(i)}(x) \cap \tilde{\Xi}\}$ by a finite number of set theoretic operations (intersections, set differences). That such a partition exists follows from the fact that $x \in K_2^{p}$ implies that

$\bigcup \Xi_{(i)}(x) \supset \tilde{\Xi}$. Now, define $\bar{\nu}_{(j)}(x,\xi)$ on $\Sigma_{(j)}(x)$ as follows

$$\bar{\nu}_{(j)}(x,\xi) = \underset{(i)}{\mathrm{Max}}\{\nu_{(i)}(x,\xi) \mid \xi \in \Xi_{(i)}(x) \cap \Sigma_{(j)}(x)\}.$$

Then

$$Q(x) = \int_{\tilde{\Xi}} Q(x,\xi)\,d\mu = \sum_{(j)} \int_{\Sigma_{(j)}(x)} Q(x,\xi)\,d\mu \leq \sum_{(j)} \int_{\Sigma_{(j)}(x)} \bar{\nu}_{(j)}(x,\xi)\,d\mu.$$

The last member of this inequality is a finite sum of finite terms, from which it follows that $Q(x)$ is bounded above. Thus $K_2^s \supset K_2^p$, which completes the proof.

The first equality in (4.2) extends to the case where W is also random [Walkup and Wets, 1967-b] provided one requires that the restriction of pos $W(\xi)$ to $\tilde{\Xi}$ is continuous, where pos $W(\xi)$ is viewed as a map from $\Xi \subset R^{N+(\bar{n}\times\bar{m})}$ into a metric space $(\mathcal{C},d)$ whose points are convex cones with apex at the origin and the metric d is determined by the Hausdorff distance between their intersections with the unit ball [Walkup and Wets, 1967-a]. However, when W is random, there seems to be no easily verifiable condition which one could impose on Ξ so as to insure that the second equality in (4.2) also remains valid.

From the definition of the problem (2.2), it is easy to see that feasibility depends only on the p, T components of ξ. Thus if we let $\Xi_{p,T}$ be the projection of $\tilde{\Xi}$ on the p,T coordinates of R^N, we have that

$$K_2^p = \bigcap_{\xi \in \tilde{\Xi}} K_2(\xi) = \bigcap_{\xi \in \Xi_{p,T}} K_2(\xi)$$

where

(4.4) $$K_2(\xi) = \{x \mid p(\xi) - T(\xi)x \in \text{pos } W\} = \{x \mid Q(x,\xi) < +\infty\}.$$

By arguments similar to those used in the first part of the proof of theorem (4.1), one can show that $K_2 = \cap_{\xi \in \tilde{\Xi}_{p,T}} K_2(\xi)$ where $\tilde{\Xi}_{p,T}$ is the support of the marginal distribution of $(p(\xi),T(\xi))$.

Corollary (4.5). For a stochastic program with fixed recourse, where ξ satisfies a weak covariance condition, such as (2.1), the relation

$$K_2 = K_2^p = K_2^\mu = K_2^s$$

holds, where

$$K_2 = \bigcap_{\zeta \in \tilde{\Xi}_{p,T}} K_2(\zeta) = \{x \mid p(\zeta)-T(\zeta)x \in \text{pos } W \text{ for all } (p(\zeta),T(\zeta)) \text{ in } \tilde{\Xi}_{p,T}\}$$

In view of the preceding corollary, we shall write K_2 whenever we want to refer to a set of constraints induced on x by the stochastic constraints of problem (2.3): $T(\xi)x + Wy = p(\xi)$ a.e. This set K_2 possesses various interesting properties which we investigate in the remainder of this section.

Theorem (4.6). Let Σ be a set obtained from $\tilde{\Xi}_{p,T}$ by applying the operations: topological closure, convex closure, positive closure, positive scalar multiplication, or any of the (not necessarily unique) inverses of these operations, then

$$K_2 = \bigcap_{\zeta \in \Sigma} K_2(\zeta).$$

(The inverse operation of convex closure yields the extreme points.)

Proof. Suppose $K_2 = \bigcap_{\zeta \in \Sigma'} K_2(\zeta)$, then the fact that pos W is closed justifies replacing Σ' either by its closure or by a dense subset of Σ' in the definition of K_2. The remaining assertions of the theorem follow rather directly from the fact that pos W is a convex cone and the definition of $K_2(\xi)$ given by (4.4).

<u>Theorem (4.7)</u>. The feasibility region K_2 is a closed convex subset of R^n. Moreover if the closed positive hull of $\tilde{\Xi}_{p,T}$, written pos $(\tilde{\Xi}_{p,T})$, is a convex polyhedral cone, then K_2 is a convex polyhedron.

<u>Proof</u>. Since $K_2 = \bigcap_{\zeta \varepsilon \tilde{\Xi}_{p,T}} K_2(\zeta)$, the first part of the theorem follows trivially from the observation that for each ζ, $K_2(\zeta)$ is a closed convex polyhedron. The second part follows from theorem (4.6) since in this case we may replace $\tilde{\Xi}_{p,T} \subset R^{(n+1)\bar{m}}$ by a finite set, namely by any set of points whose positive combinations span $\text{pos}(\tilde{\Xi}_{p,T})$. This set of points can be chosen finite by (4.6) since by assumption $\text{pos}(\tilde{\Xi}_{p,T})$ is a polyhedron and thus has only a finite number of extremal elements. (Note that there are other cases for which K_2 is polyhedral, e.g., if pos $W = R^{\bar{m}}$, then $K_2 = R^n$.)

Note that since we are dealing with finite dimensional real vector spaces, which have the Lindelöf property, it follows that K_2 can always be represented as the intersection of at most a countable subcollection of sets $K_2(\zeta)$. Also in view of theorem (4.6), the assumption that $\text{pos}(\tilde{\Xi}_{p,T})$ is polyhedral is probably the most general type of assumption which yields an easy proof that K_2 is polyhedral. In practice one might expect that very few cases will occur when $\text{pos}(\tilde{\Xi}_{p,T})$ will not be polyhedral since $\text{pos}(\tilde{\Xi}_{p,T})$ is polyhedral if the components of p and T are independent, if p and T have a finite discrete distribution, if the convex hull of $\tilde{\Xi}_{p,T}$ is polyhedral, etc. There is however one other interesting case in which one is able to prove that K_2 is polyhedral.

Before we consider this, we introduce a concept very useful in the theory of convex polyhedra in general and stochastic programming theory in particular. We first give a precise meaning to the notion of extremal elements of a convex cone.

Definition (4.8). Let A be a matrix of row size $\bar{m}$. The columns of A can be thought of as points in $R^{\bar{m}}$. Then a subset of the columns of A, say T, constitutes a *frame* for the positive hull of A if pos A = pos T and T is minimal, i.e., for every column T^k of T we have that pos $(T \sim T^k) \neq$ pos T.

Simple examples will show that a set of points in $R^{\bar{m}}$ might contain more than one frame and that two different frames do not necessarily have the same cardinality. Naturally, every finite set of points in $R^{\bar{m}}$ contains a finite frame.

Definition (4.9). The matrix A* is a *polar matrix* of A if

(i) pos A = $\{t \mid A^*t \leq 0\}$

(ii) The rows of A* constitute a frame of the polar cone pos* A of pos A, where pos* A = $\{\tau \mid \tau t \leq 0$ for all $t \in$ pos A$\}$.

Similarly, one could define the polar matrix of A as a matrix whose rows determine a set of minimal supports for the convex polyhedral cone generated by the points represented by the columns of A. In some sense, the polar matrix is a *positive* inverse of W. The notion of polar matrix was first introduced in [Wets, 1966-c] in connection with the first proof of theorem (4.10) below. A similar concept, but used in a different context, can also be found in the work of Fulkerson [Fulkerson, 1968].

He uses the name of *blocking matrix*.

In general, the number of rows of A^* is not commensurate with the numbers of columns of A. In particular A^* might have no rows, which corresponds to the case when pos A has no supports, i.e., when pos $A = \bar{R}^m$. On the other hand, if pos A is a cone over a neighborly polytope [Grünbaum, 1967; Chap. 7], then pos A might contain an extremely large number of rows when compared with the number of columns of A. Various efforts have been made to relate the generalized inverse of a matrix A to the polar matrix of A. So far none of these attempts have been remunerated by success. One might reasonably expect that no interesting relation does exist, [Wets, 1968].

<u>Theorem (4.10)</u>. Suppose T is fixed in the stochastic program (2.2) and ξ satisfies a weak covariance condition, then K_2 is a closed convex polyhedron.

<u>Proof</u>. Since T is fixed, by corollary (4.5) $x \in K_2$ if and only if $(p(\xi)-Tx) \in \text{pos } W$ for all $p(\xi)$ in $\tilde{\Xi}_p$ where $\tilde{\Xi}_p$ is the support of the distribution of $p(\xi)$. Now $(p(\xi)-Tx) \in \text{pos } W$ if and only if $W^*(p(\xi)-Tx) \le 0$. Thus $x \in K_2$ if and only if $(W^*T)x \ge W^*p(\xi)$ for all $p(\xi)$ in $\tilde{\Xi}_p$ or, by theorem (4.6), for all $p(\xi)$ in Σ where Σ is the closed convex hull of $\tilde{\Xi}_p$. This possibly infinite system of linear inequalities can be replaced by

$$(4.11) \qquad (W^*T)_i x \ge \alpha_i^* = \sup_{p(\xi)\in\Sigma} W_i^* p(\xi) \qquad i = 1,\dots,\ell$$

where W_i^* denotes the i^{th} row of the matrix W^* and ℓ is the number of rows of the matrix W^*. Unless $\alpha_i^* < +\infty$ for all $i = 1,\ldots,\ell$ the problem is infeasible in which case $K = \emptyset$ and the theorem holds trivially, otherwise the system (4.11) constitutes a finite system of linear constraints which determine the polyhedron

$$(4.12) \qquad K_2 = \{x \mid (W^*T)x \geq \alpha^*\}$$

where α^* is the vector whose components are α_i^*.

Corollary (4.13). If $p(\xi)$ and $T(\xi)$ are independent and $\tilde{\Xi}_T$ (or the closure of its positive hull) is polyhedral, then K_2 is polyhedral.

Proof. Since $p(\xi)$ and $T(\xi)$ are independent $\tilde{\Xi}_{p,T} = \tilde{\Xi}_p \times \tilde{\Xi}_T$. For each $T(\xi)$ in $\tilde{\Xi}_T$, let $K_2(T(\xi))$ denote the set of x's such that

$$[W^*T(\xi)]x \geq \alpha^*$$

Now $K_2 = \bigcap_{\zeta\in\tilde{\Xi}_{p,T}} K_2(\zeta) = \bigcap_{T(\xi)\in\tilde{\Xi}_T} K_2(T(\xi))$. Since $\tilde{\Xi}_T$ is polyhedral, by theorem (4.6) we may replace $\tilde{\Xi}_T$ in $\bigcap_{T(\xi)\in\tilde{\Xi}_T}$ by a finite set representing the extremal elements of $\tilde{\Xi}_T$. Thus K_2 can be written as a finite intersection of polyhedra $K_2(T(\xi))$.

So far we have shown that the region of feasibility of a stochastic program with fixed recourse, i.e., the set

$$K = K_1 \cap K_2$$

is a set possessing fairly interesting properties. It is always closed and convex and under fairly general assumptions it is even polyhedral. When T is fixed, we have also seen that $x \in K$ if and only if x satisfies some deterministic linear constraints (4.12) where α^* represents in some sense a lower bound of the set $\tilde{\Xi}_T$. Determining such a lower bound--especially if it is easy to compute--might be very useful in obtaining criteria of feasibility as well as for constructing algorithmic procedures generating feasible solutions. In this pursuit it is convenient to introduce some terminology related to convex cone ordering of real vector spaces. This approach was first presented in some lectures given at the University of California [Wets, 1967]. In his doctoral dissertation [Parrikh, 1967] Parrikh shows that these ideas can also be fruitfully exploited in a slightly different setting.

<u>Definition (4.14)</u>. The partial ordering $\precsim$ is said to be a cone ordering induced by a closed convex cone $C \subset \bar{R}^m$ if

$$x \precsim y \text{ is equivalent to } y - x \in C$$

We need the following obvious property of cone orderings:

<u>Proposition (4.15)</u>. Let $\precsim_{C_1}$ denote the cone ordering induced by the closed convex cone C_1 and let $\precsim_{C_2}$ denote the cone ordering induced by the closed convex cone C_2. Then

$$x \precsim_{C_1} y \precsim_{C_2} z \text{ implies } x \precsim_{C_1+C_2} z$$

where $C_1 + C_2$ denotes the vector sum of C_1 and C_2. In particular if $C_1 \supset C_2$ then

$$x \precsim_{C_1} y \precsim_{C_2} z \text{ implies } x \precsim_{C_1} z$$

<u>Definitions (4.16)</u>. Let Σ be a subset of $\overline{R}^m$ and $\lesssim_C$ a cone ordering induced by C. A point α_C is said to be a *greatest lower bound* of Σ with respect to the ordering induced by the cone C if $\alpha \lesssim_C \alpha_C$ for all points α satisfying $\alpha \lesssim_C \zeta$ for all ζ in Σ.

Moreover if α_C also belongs to the closure of Σ, then α_C is said to be a *proper lower bound* of Σ.

<u>Theorem (4.17)</u>. Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Let C be a closed convex cone contained in pos W, $\Delta(x) = \{p(\xi) - T(\xi)x \mid p(\xi), T(\xi) \in \tilde{\Xi}_{p,T}\}$ and $\alpha_C(x)$ is a proper lower bound of $\Delta(x)$. Then $x \in K_2$ if and only if $\alpha_C(x) \in$ pos W.

<u>Proof</u>. By theorem (4.6), $x \in K_2$ if and only if the closed set $\Delta(x) \subset$ pos W. The rest of the proof follows from the last part of proposition (4.15) since by assumption $C \subset$ pos W.

<u>Corollary (4.18)</u>. Suppose T is fixed in a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Suppose C is a closed convex cone contained in pos W and α_C is a proper lower bound of $\tilde{\Xi}_p$. Then $x \in K_2$ if and only if $\alpha_C(x) = \alpha_C - Tx \in$ pos W. In particular, $x \in K_2$ if and only if $\alpha_{\text{pos } W} - Tx \in$ pos W, where $\alpha_{\text{pos } W}$ is a proper lower bound of $\tilde{\Xi}_p$ with respect to $\lesssim_{\text{pos } W}$. Moreover, $\alpha_{\text{pos } W} = \alpha^*$ where α^* is as defined by (4.11).

Proof. The first part of the corollary is an immediate consequence of the theorem and the equality $\alpha^* = \alpha_{\text{pos } W}$ follows directly from their respective definitions.

Provided that a proper lower bound exists, the usefulness of the above characterizations is only limited by our capability of computing it. When T is fixed, one way to determine this proper lower bound is to compute W^*. This might be by itself a major undertaking which we shall discuss further in the following section. In practice, finding such a lower bound reduces down to finding some cone C contained in pos W such that the proper lower bound, with respect to the ordering induced by C, of any subset in $\bar{R}^m$ is fairly easy to compute. This will certainly be the case if C can be selected to be some orthant. If a lower bound of $\tilde{\Xi}_{p,T}$ with respect to some orthant is also a proper lower bound, then by theorem (4.17) and its corollary, verifying feasibility of any given x is a fairly easy task. We shall devote a fair portion of the subsequent developments to the case when pos W contains some orthant. Moreover, the following proposition helps justify the relative importance we attach to this case.

Proposition (4.19). Let $\lesssim_C$ be a cone ordering on $\bar{R}^m$. Then every bounded subset of $\bar{R}^m$ has a unique greatest lower bound if and only if the cone C is the positive orthant with respect to some coordinate system.

Proof. The if condition is obvious. To prove the only if conditions we make use of Choquet's characterizations of simplices in terms of homothetic intersections. Suppose every bounded subset of $R^{\bar{m}}$ has a unique greatest lower bound with respect to $\leqslant_C$. It is easy to verify that uniqueness implies that C must be a pointed cone and that the existence of lower bounds for full dimensional subsets of $R^{\bar{m}}$ implies that C has dimension $\bar{m}$. Let H be some hyperplane supporting C at the origin only, H_+ the closed half-space bounded by H which contains C, H_+^c the complement of H_+ and σ in H_+^c the unit normal to H. Let s and t be any two distinct points of H_+^c, then $P_s = (s+C) \cap H_+^c$ and $P_t = (t+C) \cap H_+^c$ are homothetic, with s and t greatest lower bounds for P_s and P_t respectively. Suppose $Q = P_s \cap P_t \neq \emptyset$. By the hypotheses Q has a greatest lower bound q and thus $(q+C) \cap H_+^c = P_q \supset Q$. Since $P_s \supset Q$, s is a lower bound for Q and $s \leqslant_C q$. Thus $P_s \supset P_q$ and similarly $P_t \supset P_q$ so that $P_q \supset Q = P_t \cap P_s \supset P_q$, i.e., $P_s \cap P_t = P_q$. Thus any homothets of the closed bounded figure P_σ intersect either in a point, a homothet of P or the empty set. By Choquet's characterization [Choquet, 1956]--for a simple proof see [Eggleston, Grünbaum and Klee, 1964]--P_σ must be a simplex. Since C has full dimension, it follows that C is a simplicial cone which is linearly isomorphic to the positive orthant of $R^{\bar{m}}$.

Theorems (4.7), (4.10) and (4.17) and their corollaries summarize the most useful characterization of the feasibility region K_2 which are so far available. Among other things they allow us to rewrite the deterministic equivalent program as

$$\text{(4.20)} \qquad \begin{array}{ll} \text{Find inf } z(x) = \bar{c}x + Q(x) & \\ \text{subject to} \quad Ax & = b \\ (W^*T(\zeta))x \geq W^*p(\zeta) & \text{for all } \zeta \text{ in } \tilde{\Xi}_{p,T} \\ x \geq 0 & \end{array}$$

where the function $Q(x)$ is finite on the set determined by the constraints unless, as we shall see, it is identically $-\infty$ on this set. If the structure of $\tilde{\Xi}_{p,T}$ allows us to use some of the preceding results, then even more practically oriented expressions can be found for the deterministic equivalent problem. For example, if T is fixed, then one can write (4.20) as

$$\text{(4.21)} \qquad \begin{array}{ll} \text{Find inf } z(x) = \bar{c}x + Q(x) & \\ Ax & = b \\ (W^*T)x & \geq \alpha^* \\ x & = 0 \end{array}$$

where α^* is defined by (4.11).

5. <u>Finding a Feasible Solution</u>.[†] In the first part of this section, we outline some very general ideas which can be used to find a feasible solution. These necessary concepts are summarized in the algorithm (5.5) and its variants described below. So much will depend on the particular structure or the particular way the entities involved are given that in fact it is impossible to find a unique "best" procedure. The algorithms that we describe should be viewed as possible skeletons rather than finished products. In the second part of this section we deal with some specific cases which allow for very effective solution procedures.

[†] The results derived in the first part of this section were obtained in collaboration with D. Walkup.

Determining whether a point x belongs to K_1 reduces to checking if it satisfies a system of linear inequalities. We consider this problem solved and thus limit ourselves to the case of finding a point of K_2. In the preceding section we have obtained various characterizations of the set K_2. In general, i.e., when the support of the distribution or the matrix W have no special properties, determining if a given x is in K_2 is equivalent to determining if

$$\Delta(x) = \{t \mid t = p(\xi) - T(\xi)x, \quad (p(\xi),T(\xi)) \in \Sigma\}$$

is or is not contained in the convex polyhedral cone pos W, where Σ is a set obtained from $\tilde{\Xi}_{p,T}$ by any one of the operations mentioned in theorem (4.6). If Σ is a convex set (polyhedron) so is $\Delta(x)$. Thus in general determining if a given x is feasible comprehends the problem of determining when a convex set is contained in a convex polyhedral cone. Depending on the manner in which the set $\Delta(x)$ is given, this problem might turn out to be fairly easy or extremely difficult. If the set $\Delta(x)$ is finite or more generally if its convex hull has only a finite number of extreme elements (extreme rays and extreme points) [Klee, 1957] which can be easily determined, then it is not very difficult to determine if $\Delta(x) \subset \text{pos } W$, since this involves at most solving a finite number of linear programs which differ only in their right-hand sides, namely,

$$\text{(5.1)} \qquad \text{Minimize } w = es^+ + es^-$$
$$Wy + Is^+ - Is^- = \delta^k_{(x)}$$
$$y \geq 0, \; s^+ \geq 0, \; s^- \geq 0$$

where e is a row vector of size $\bar{m}$ whose components are 1's and $\delta^k_{(x)}$, $k = 1,\ldots,t$ are the extremal elements of the convex hull of $\Delta(x)$. Various tricks are available which allow for considerable simplification and the work involved is by no means to be equated with solving t linear programs [Van Slyke and Wets, 1969, Section 5]. Although in general finding the extremal elements of $\Delta(x)$ might prove to be a real challenge, in some particular cases this might not prove to be very difficult, e.g., if the components of $T(\xi)$ and $p(\xi)$ are independent, then the extremal elements of $\Delta(x)$ can be obtained by paying attention only to the extremal elements of $\tilde{\Xi}_{p,T}$ which in this case is a rectangle (possibly unbounded) in $R^{\bar{m}(n+1)}$. If either $\Delta(x)$ (or its convex hull) is only available in terms of the bounding hyperplanes of the set or when $\Delta(x)$ is not even polyhedral, then one would have to resort to a technique somewhat similar to the one described below.

Feasibility Criterion (5.2). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. A vector x in K_1 is a feasible solution if and only if the optimal value of each of the ℓ convex programs

$$(5.3) \qquad \text{Find } \inf_{(p,T)} \; w = W^*_k(Tx-p)$$

$$\text{subject to } (p,T) \in \Sigma$$

is greater than or equal to zero, where W^*_k is the k^{th} row of W^* the polar matrix of W, ℓ is the number of rows of W^* and Σ is the closed convex hull of $\tilde{\Xi}_{p,T}$.

This feasibility criterion is nothing more than a reformulation of the induced constraints found in the deterministic equivalent program (4.20). The following proposition provides a method for "improving" an infeasible solution. In fact, the proposition shows how one can generate bounding hyperplanes of the set K_2.

Proposition (5.4). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Suppose Σ is a set obtained from $\tilde{\Xi}_{p,T}$ by topological, convex or positive closure and $\hat{x}$ is any point in R^n. If (p,T) is a point of Σ such that $W_k^*(T\hat{x}-p) < 0$ then either

$$\{x \mid (W_k^* T)x \geq W_k^* p\}$$

is empty or is a closed halfspace in R^n containing K_2 but not $\hat{x}$. If (p,T) and (p',T') are points in $R^{\bar{m}(n+1)}$ such that for all sufficiently large positive λ, $(p^\lambda,T^\lambda) = (p,T) + \lambda(p',T')$ is a member of Σ and $W_k^*(T^\lambda\hat{x}-p^\lambda)$ is unbounded below as $\lambda \to +\infty$, then either

$$\{x \mid (W_k^* T')x \geq W_k^* p'\}$$

is empty or a closed halfspace in R^n containing K_2 but not $\hat{x}$.

Proof. The first part follows directly from the feasibility criterion (5.2). For the second part if $W_k^*(T^\lambda\hat{x}-p^\lambda) = W_k^*(Tx-p) + \lambda W_k^*(T'x-p')$ goes to $-\infty$ as λ goes to $+\infty$, then $W_k^*(T'\hat{x}-p')$ must be strictly negative. Combining this last observation with the feasibility criterion (5.2) completes the proof.

This last proposition suggests the following algorithm for producing a feasible solution $\bar{x}$ in K.

Algorithm (5.5).

(i) At the start of the i^{th} iteration a set S^i of linear equations and inequalities in the variables x are given. (At the start of the first iteration S^1 is just the set of linear relations determining K_1, viz. $Ax = b$, $x \geq 0$.)

(ii) A feasible solution x^i satisfying the linear relations S^i is sought. If none exists the stochastic program is declared infeasible and the algorithm terminates.

(iii) For each row W_k^* of the polar matrix W^* in turn, the convex program (5.3) is solved with $x = x^i$. If all ℓ programs have nonnegative values, x^i is defined to be $\bar{x}$ and the algorithm terminates. Otherwise for some k a point (p,T) or points (p,T) and (p',T') as in proposition (5.4) are found. In this case the appropriate inequality $(W_k^*T)x \geq W_k^*p$ or $(W_k^*T')x \geq W_k^*p'$ is added to S^i to form S^{i+1} and the $(i+1)^{st}$ iteration is started from step (i).

If Σ is a convex polyhedron given by linear equations and linear inequalities, the programs (5.3) to be solved during step (iii) of the above algorithm are linear programs. We can thus be more explicit in the description of step (iii):

(iii') ... Otherwise for some k one finds either an optimal basic solution (p,T) yielding negative objective or a basic solution and a direction (p',T') corresponding to an unbounded feasible pivot from (p,T)...

In this case we can prove:

Proposition (5.6). Suppose Σ used in the definition of K_2 can be selected to be a convex polyhedron defined by linear relations. Then algorithm (5.5) using step (iii') terminates in a finite number of steps. It either generates a feasible solution of the program or it establishes that the problem is infeasible.

Proof. It suffices to show the finiteness of the process. This follows immediately from the fact that there are only a finite number of rows in W^* and only a finite number of basic solutions and unbounded basic pivots for each problem (5.3). Thus there are only a finite number of inequalities which can be added to S^1. Moreover, once an inequality has been added it cannot be generated again.

Note that in fact this proposition gives another proof of the polyhedral property of K_2 under the hypotheses of the second part of theorem (4.7). If Σ is not polyhedral, there does not seem to be any condition one could impose on the selection of (p,T) or (p',T') in step (iii) of the algorithm (5.5) which would insure that the algorithm converges, unless, perhaps, one is satisfied with an epsilon type of feasibility as in [Parrikh, 1967].

Not even taking into account the substantial amount of work which the computation of W^* might necessitate, the somewhat laborious fashion by which the algorithm (5.5) generates a feasible solution is due to the fact that to test the feasibility of any given x, one has to take into consideration if not all points of $\tilde{\Xi}_{p,T}$ (or some set derived from it) at least a sufficiently large number so as to make the process inefficient. In the previous section, we have seen that a considerable simplification is possible when the sets $\Delta(x)$ have a proper lower bound $\alpha_C(x)$ with respect to the ordering induced by some convex cone C contained in pos W. If this is the case, it is no longer necessary to consider the whole set $\tilde{\Xi}_{p,T}$, or some large subset of it to verify the feasibility of a given x. As shown by theorem (4.17) it will be sufficient to solve only *one* linear program of the form (5.1), namely,

$$
\text{(5.7)} \qquad \begin{aligned} \text{Minimize } w = \quad & es^+ + es^- \\ & Wy + Is^+ - Is^- = \alpha_C(x) \\ & y \geq 0,\ s^+ \geq 0,\ s^- \geq 0 \end{aligned}
$$

This enormous simplification naturally depends on our capability of finding a convex cone C in pos W, such that the proper lower bound of any set will be fairly easy to compute. In the remainder of this section we shall assume that the positive orthant pos I (or some other orthant which can always be made to be the positive orthant by an appropriate change in sign of some of the rows of the equations $T(\xi)x + Wy = p(\xi)$) is contained in pos W. Given the practical use we want to make of this assumption, proposition (4.19) justifies the restriction to this case.

If the convex cone C of theorem (4.17) and its corollary can be selected to be the positive orthant, then $\Delta(x)$ has a proper lower bound if it contains a point α such that $\alpha_i \leq \delta_i$, $i = 1,\ldots,\bar{m}$ for all δ in $\Delta(x) \subset R^{\bar{m}}$. This ordering induced by pos I is usually referred to as the componentwise ordering. Thus in this case verifying if $\Delta(x)$ has a proper lower bound, and if it does, finding it, is usually fairly easy to do. In particular if T is fixed, then

$$\Delta(x) = \{p(\xi)\,|\,p(\xi) \in \tilde{\Xi}_p\} - Tx$$

and as shown in corollary (4.18) it suffices to find a proper lower bound of $\tilde{\Xi}_p$ (or its closed convex hull) with respect to the componentwise ordering to determine a lower bound of $\Delta(x)$ for all x. In this case it is determined by $\alpha_{\text{pos I}} - Tx$, where $(\alpha_{\text{pos I}})_i \leq p_i$ for all p in $\tilde{\Xi}_p$ (or its convex hull).

Feasibility Criterion (5.8). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Suppose that for all x the set $\Delta(x)$ possesses a proper lower bound $\alpha_C(x)$ with respect to the ordering induced by a closed convex cone C contained in pos W. Then $\hat{x}$ in K_1 is feasible if and only if the system of linear relations

$$\text{(5.9)} \qquad \sigma W \leq 0, \quad \sigma\alpha_C(\hat{x}) > 0$$

is inconsistent, where σ is an $\bar{m}$-row vector of variables σ_i. In particular, if $C = \text{pos I}$ and T is fixed, then $\hat{x}$ in K_1 is feasible if and only if the system of linear relations

$$\text{(5.10)} \qquad \sigma W \leq 0, \quad \sigma(\alpha - T\hat{x}) > 0$$

is inconsistent, where α is a proper lower bound of $\tilde{\Xi}_p$

with respect to the componentwise ordering, i.e., $\alpha_i \leq p_i$ for all p in $\tilde{\Xi}_p$.

Proof. This feasibility criterion is an immediate consequence of theorem (4.17) and its corollary, since if either system above is solvable, then there exists a hyperplane, determined by its normal σ, separating pos W and $\alpha_C(\hat{x})$ in the first case and pos W and $\alpha - Tx = \alpha_{\text{pos } I}(\hat{x})$ in the second case.

For computational purposes, the following proposition supplements the feasibility criterion given above.

Proposition (5.11). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition and suppose that pos W contains the positive orthant pos I. Suppose further that the rows of T and p are independent and that each $\tilde{\Xi}_i$, the support of the random variables in row i of T and p, is bounded. Then $\alpha_{\text{pos } I}(x)$ exists for all x and is given by

$$\alpha_i(x) = \text{Min}\{(p_i - T_i x) \mid (p, T) \in \tilde{\Xi}_i\}$$

where $\alpha_i(x)$ is the i^{th} component of $\alpha_{\text{pos } I}(x)$. Further if the columns of T and p are also independent, then

$$\alpha_i(x) = \underset{\tilde{\Xi}p_i}{\text{Min}}\, p_i - \sum_{j=1}^{n} \underset{\tilde{\Xi}t_{ij}}{\text{Max}}\, t_{ij} x_j$$

The feasibility criterion (5.8) is a generalization of the feasibility test found in [Wets, 1966-b; Section 2]. It is not difficult to see that if $\hat{x}$ is infeasible and we find some σ satisfying (5.9) or (5.10), then we are in a very similar situation to that in algorithm (5.5) when we found some row W_k^* of the polar matrix of W and a point (p,T) of Σ which could be used to generate a linear constraint which has to be satisfied by every feasible x but which fails to be satisfied by $\hat{x}$. If $\alpha_C(\hat{x}) = p_C - T_C\hat{x}$, where (p_C,T_C) belongs to $\tilde{\Xi}_{p,T}$, then if σ satisfies (5.9), the hyperplane

$$(\sigma T_C)x = \sigma p_C$$

separates K_2 from $\hat{x}$. When T is fixed the closed halfspace containing K_2 determined by this hyperplane, can be written as

$$(\sigma T)x \geq \sigma\alpha. \tag{5.12}$$

The feasibility constraints of type (5.12) bear more than a passing resemblance to those found in proposition (5.4).

In the remaining part of this section we limit ourselves to the case when T is fixed. The extension of the results below to the case when T also contains random elements is essentially routine, but requires the introduction of more than a reasonable amount of cumbersome notation which seem hardly justified in this survey. The algorithm, in this case, is again a modification of the algorithm (5.5). We replace step (iii) by

(iii'') Solve the linear program

(5.13) Minimize $w = es^+ + es^-$

$$Wy + Is^+ - Is^- = \alpha - Tx^i$$

$$y \geq 0,\ s^+ \geq 0,\ s^- \geq 0$$

where α is a lower bound of $\tilde{\Xi}_p$ with respect to the ordering induced by pos I. Since pos $W \subset$ pos I, the artificial variables s^+ can be deleted from the program (5.13) without impairing its feasibility. However, if we incorporate them in our formulation, they guarantee a starting basis. If $w = 0$, then x^i is a feasible solution and the algorithm terminates. Otherwise at the optimum there is a set of optimal multipliers σ such that $\sigma(\alpha - Tx^i) > 0$ and $\sigma W \leq 0$ (compare these relations with those of the second part of the feasibility criterion (5.8)). In this case the inequality

$$(5.14) \qquad (\sigma T)x \geq \sigma\alpha$$

is added to S^i to form S^{i+1} and the $(i+1)^{st}$ iteration is started from step (i).

Proposition (5.15). Suppose T fixed in a stochastic program with fixed recourse whose random variables satisfy a weak covariance condition. Moreover, suppose that α is a proper lower bound of the closed convex hull of $\tilde{\Xi}_p$ with respect to the ordering induced by pos I and pos I $\subset$ pos W. Then, the algorithm (5.5) with step (iii'') terminates in a finite number of steps. It either generates a feasible solution or recognizes that the stochastic program is infeasible.

Proof. The matrix (W,I,-I) of program (5.13) contains only a finite number of bases. Thus only a finite number of optimal multipliers of type σ can be generated. Moreover, once a constraint of type (5.14) has been introduced in the set S^i it can never be generated again.

In order for the minimum of w in (5.13) to be positive, the optimal solution must involve some of the artificial variables (s^+, s^-) at some positive level since otherwise $w = 0$, which implies that $(\alpha - Tx^i)$ is in pos W and hence x^i in K_2. Thus in fact the number of possible σ is much smaller than *all* possible bases contained in (W,I,-I). These vectors σ are in fact normals to bounding hyperplanes of pos W which support pos W in a set containing the origin. We have that

Proposition (5.16). Suppose that the basic optimal solution to (5.13) with $w > 0$ contains exactly one of the artificial variables s^-. Then σ is the normal of a supporting hyperplane of pos W which intersects (supports) pos W in a facet, i.e., an $(\bar{m}-1)$-dimensional face of pos W.

Proof. By assumption (see Section 2) pos W is of dimension $\bar{m}$. The proof is complete if we observe that the optimal basis contains $(\bar{m}-1)$ linearly independent points of pos W contained in a hyperplane supporting pos W.

Thus if Proposition (5.16) applies, *we are generating some row of a polar matrix of* W. Moreover, every row of W* can arise in this way. However, it is not always possible to obtain the normal of a facet of pos W by solving (5.13). In [Van Slyke and Wets, 1969; Sections 2.D and 2.E] it is shown that even though this process will in some sense yeild very good constraints on x, in general [Van Slyke and Wets, 1969; Proposition (25) and Corollary (27)] it will generate more than a minimal set of supporting hyperplanes of pos W. Thus in general the row matrix obtained from the normals of these hyperplanes does not constitute a minimal set as required by the definition (4.9) of W*.

6. <u>Characterizing Relatively Complete Recourse</u>. As mentioned in the beginning of Section 4, the models studied in [Dantzig, 1955], [Madansky, 1960], [Dantzig and Madansky, 1961], [Charnes, Cooper and Thompson, 1965], and [Kall, 1967] satisfy the relatively complete recourse condition, i.e., for all x in K_1 and for all ζ in $\tilde{\Xi}_{p,T}$ there exists $y \geq 0$ such that $Wy = p(\zeta) - T(\zeta)x$. Since practical considerations seem to indicate that in applications this condition will be very often satisfied and when it is, the work involved in solving the stochastic program is considerably simplified, it is of paramount importance to be able to determine whether this condition is satisfied or not. The first efforts in this direction can be found in [Wets, 1966-b], [Kall, 1967], [Wets, 1966-c] and Dempster, 1968]. The hindsight that these studies have given us, shows that this problem has close ties [Wets, 1966-c], [Dempster, 1968, Section 5], with the theory of *positive linear dependence* developed is due by [Davis, 1954], [McKinney, 1962], [Bonnice and Klee, 1963], and more recently by [Reay, 1965-a], [Reay, 1965-b] and [Hansen and Klee, 1969].

The study of various special forms of stochastic programs with recourse [Walkup and Wets, 1969-a] suggests that it is convenient to distinguish the following cases:

<u>Definition (6.11)</u>. A stochastic program is said to have

(i) *relatively complete recourse* if $K_2 \supset K_1$;

(ii) *complete recourse* if pos $W = \bar{R}^m$;

(iii) *simple recourse* if $W = (I,-I)$, up to permutations of rows and columns if necessary.

Obviously, if $W = (I,-I)$ then $\text{pos } W = \bar{R}^m$ and thus simple recourse implies complete recourse which in turn implies relatively complete recourse. This terminology was introduced in [Walkup and Wets, 1969-a]. (It differs from the one used in [Wets, 1966-a] where the problem with simple recourse was designated as the complete problem in recognition of the fact that it is a special case of (6.1.ii). Subsequent development of the subject has suggested the distinction given here). To verify if a given problem has relatively complete recourse theoretically involves computing K_2 and then verifying if the convex polyhedral set K_1 is contained in the convex set K_2. Thus, once more we encounter here the problem of determining if a convex polyhedron is contained in a convex set. As already mentioned in Section 5, this problem might be easy or difficult, all depending on the manner in which these two sets are defined. We shall not pursue this matter any further here, at least not at this level of generality. So far no general method to solve this particular problem has been investigated. In the literature we know of only one example of a problem with relatively complete recourse but not complete recourse. Such a model was formulated by Tintner [Tintner, 1960] in connection with the allocation of available resources in agriculture economics. This model was later given the name of *active approach* to stochastic programming [Sengupta, Tintner, and Millham, 1963]. In [Walkup and Wets, 1969-a, Section 4] it is shown how this problem can be approached advantageously by the techniques developed for stochastic programs with recourse.

Rather than seeking to characterize relatively complete recourse, the research has concentrated on the characterization of complete recourse.

In [Kall, 1967, Section 4] one finds various theorems yielding sufficient conditions for complete recourse. These theorems correspond essentially to matrix versions of the caratheodory and Stinitz theorems [Reay, 1965-a] for polyhedral cones. However, as mentioned in the beginning of this section, this problem can best be handled in a somewhat more general setting. In [Wets, 1966-c] and perhaps even more clearly in [Dempster, 1968, Section 5] one can see that the basic question can be formulated as: Given an $m \times n$ matrix A determine the (facial) structure of the cone pos A generated positively by the points corresponding to the columns of A; in particular, determine if A contains a subset of columns which constitute a positive basis for R^m.

Definition (6.2). A subset of columns $(A^1,\ldots,A^k) = B$ of the matrix A is a *positive basis* for R^m, if pos $B = R^m$ and the columns of B constitute a frame (4.8), i.e., are positively linearly independent.

In the articles devoted to the theory of linear dependence, various properties of positive bases have been found. One particularly useful characterization can be found in [Reay, 1965-a]. An algebraic version of his theorem can be found in [Wets and Witzgall, 1968, Proposition 9]. An algorithm determining if a given cone pos A does or does not contain a positive basis can be found in [Wets and Witzgall, 1967]. The two last references deal with more general problems related to the algebraic characterization of the facial structure of convex polyhedral cones. As far as we are concerned here, one can find there some indication of the work involved in obtaining frames, k-faces for pos W and in particular the polar matrix of W. The particular problem of determining if a stochastic program has complete recourse can be settled by finding the lineality space of pos W [Wets and Witzgall, 1967, Section 4], where the *lineality space* denoted by $\mathcal{L}$ pos W is defined as the union of all the lines contained in pos W. In [Wets and Witzgall, 1967] it can be seen that finding $\mathcal{L}$ pos W amounts to a fairly small amount of computational work. Obviously, we have the following proposition:

Proposition (6.3). A stochastic program has complete recourse if and only if the dimension of $\mathcal{L}$ pos W is $\bar{m}$.

In general, if $\bar{m} - \dim(\mathcal{L}\text{pos } W)$ is small, the effort involved in finding the lineality space is by no means wasted, since in this case computing the polar matrix W^* would be very easy. This would allow us to use some of the algorithmic procedures described in the previous section which involve as prerequisite the computation of W^*.

7. The Objective Function. The three preceding sections have been essentially devoted to obtaining various properties of the set K_2 on which $Q(x) < +\infty$, which with K_1 determine the feasibility region of the stochastic program (2.2). In this section we shall be especially concerned with the properties of $z(x)$ on K or more particularly of $Q(x)$ on K_2. In order to derive these properties we shall rely on some results from the perturbation theory for linear programming. Although we could use some more elementary facts to obtain certain of the desired results, the theorem (7.2) below is probably best suited to our purposes, and is also a very useful conceptual tool in the general area of stochastic programming. It has been used [Walkup and Wets, 1968] to study the properties of decision rules, so

prominent in literature devoted to stochastic programs with chance-constraints. In order to allow us to rely on the intuitive geometric content of this result, we need the following definition.

Definition (7.1). A *finite closed polyhedral complex* will be any finite collection $\mathcal{K}$ of closed convex polyhedra, called *cells* of $\mathcal{K}$, such that:

(i) If C is a cell of $\mathcal{K}$ then every closed face of C is a member of $\mathcal{K}$.

(ii) If C_1 and C_2 are distinct cells of $\mathcal{K}$, then either they are disjoint, or one is a face of the other, or their intersection is a face of each.

Theorem (7.2) (BASIS DECOMPOSITION THEOREM). Let $P(t)$ denote the linear program

$$\begin{aligned} \text{Minimize} \quad & cx \\ \text{subject to} \quad & Ax = t \\ & x \geq 0 \end{aligned}$$

where c is fixed and A is a fixed $m \times n$ matrix of rank m. Then:

(i) $P(t)$ is feasible if and only if t lies in pos A.

(ii) Either $P(t)$ is bounded for all t in pos A or $P(t)$ is unbounded for all t in pos A.

(iii) If $P(t)$ is bounded there exists a decomposition of pos A into a finite closed polyhedral complex $\mathcal{K}$ whose cells are simplicial cones with vertex at the origin, and a one-to-one correspondence between the one-dimensional cells of $\mathcal{K}$ and selected columns of A which generate them such that

(a) the closed m-dimensional cells of $\mathcal{K}$ cover pos A, and

(b) the m columns of A associated with the edges of a closed m-dimensional cell C of $\mathcal{K}$ constitute an optimal basis for all t in C.

This theorem which is proved in [Walkup and Wets, 1969-b] has various important consequences, e.g., it follows from part (iii) that provided $P(t)$ is bounded there exists a piecewise linear continuous function $x(t)$ which determines a basic optimal solution for t in pos A. The particular consequences of interest here are given in the two following corollaries.

<u>Corollary (7.3)</u>. The function $Q(t) = \{\text{Min } cx \mid Ax = t,\ x \geq 0\}$ is a finite convex polyhedral function on pos A unless $Q(t) = -\infty$ for some t in pos A, in which case $Q(t)$ is identically $-\infty$ on pos A.

<u>Proof</u>. The finiteness and unboundedness situations are taken care of by part (ii) of the above theorem. For the remainder it suffices to observe that $Q(t)$ will be linear on every simplicial conical cell of the polyhedral complex $\mathcal{K}$ generated by the decomposition of pos A, and that the convexity of $Q(t)$ follows from the fact that if x^0 and x^1 are optimal solutions for $t = t^0, t^1$, then $(1-\lambda)x^0 + \lambda x^1$ is a feasible but not necessarily optimal solution when $t = (1-\lambda)t^0 + \lambda t^1$.

<u>Corollary (7.4)</u>. The function $Q^*(t) = \{\text{Min } tx \mid Ax = b,\ x \geq 0\}$ is a finite concave polyhedral function which for every vector t^T in $\text{pos}(A^T,-A^T,I)$ $= \{t \mid t = uA - vA + sI,\ u \geq 0,\ v \geq 0,\ s \geq 0\}^T$ unless $Q^*(t) = +\infty$ for some t^T in $\text{pos}(A^T,-A^T,I)$ in which case $Q^*(t)$ is identically $+\infty$ on

$\text{pos}(A^T,-A^T,I)$.

Proof. This corollary follows trivially from the previous corollary and a straightforward application of the standard duality theorem of linear programming.

It is now easy to see that:

Proposition (7.5). The function $Q(x,\xi) = Q(x,[q(\xi),p(\xi),T(\xi)])$ is a convex polyhedral function in x on K_2 for each ξ in $\tilde{\Xi}$. Moreover, it is concave polyhedral in $q(\xi)$ and convex polyhedral in $(p(\xi),T(\xi))$.

Proof. By definition $Q(x,\xi)$ equals $\{\text{Min } q(\xi)y \mid Wy = p(\xi) - T(\xi)x,\ y \geq 0\}$. The right-hand sides of the constraints of this problem are linear in x and $(p(\xi),T(\xi))$. Application of Corollary (7.3) yields the assertions of the proposition with respect to x and $(p(\xi),T(\xi))$. The remainder follows from corollary (7.4).

Note that the above proposition is not restricted to the domain of finiteness of $Q(x,\xi)$ but in fact holds for all x in R^n and ξ in $\Xi \subset R^N$ provided one adopts the standard convention of setting $Q(x,\xi) = +\infty$ if the constraints define an empty set and $Q(x,\xi) = -\infty$ if the problem is unbounded. This fact would allow us to prove the first assertion of the theorem below without any restriction whatsoever on the distribution of ξ; in fact, the convexity of $z(x)$ holds even when W is also a random matrix [Walkup and Wets, 1967-b; Theorem (4.4)].

Theorem (7.6). Consider a stochastic program with fixed recourse (2.2) whose random elements satisfy a weak covariance condition. Then

$z(x) = \bar{c}x + Q(x) = E_{\xi}\{c(\xi)x + Q(x,\xi)\}$ is a convex function on K. Moreover, $z(x)$ is either finite on K or $z(x)$ is identically $-\infty$ on K.

Proof. Since $\bar{c}x$ is linear in x and $K_2 \supset K$, it is sufficient to prove the above assertions with $z(x)$ and K replaced by $Q(x)$ and K_2 respectively. The convexity of $Q(x)$ follows directly from the isotone and subadditivity of the integral $\int \cdot d\mu$ and proposition (7.5) which yields the convexity of $Q(x,\xi)$ in x. By theorem (4.1) the function $Q(x)$ is less than $+\infty$ for all x in K_2, thus to complete the proof it suffices to show that if $Q(\hat{x}) = -\infty$ for some $\hat{x}$ in K_2, then $Q(x) \equiv -\infty$ for all x in K_2. Suppose $Q(\hat{x}) = -\infty$ for some $\hat{x}$ in K_2 then the set $\{\xi | Q(\hat{x},\xi) = -\infty\}$ must have positive measure. This follows from our definition of the integral and the weak covariance condition. By corollary (7.3), for any ξ in $\tilde{\Xi}$, $Q(\hat{x},\xi) = -\infty$ implies $Q(x,\xi) = -\infty$ for all x in K_2. Thus for all x in K_2 the set $\{\xi | Q(x,\xi) = -\infty\}$ has positive measure, i.e., $Q(x) = -\infty$ for all x in K_2.

In general $Q(x)$, and thus $z(x)$, are not continuous on K, but under very general conditions one can prove that $Q(x)$ is lower semi-continuous on the set on which it is finite [Walkup and Wets, 1969-c]. The lower semicontinuity of $z(x)$ is in fact sufficient to imply its continuity if we can show that K_2 is polyhedral (see Section 4), since every convex function is upper semicontinuous on a convex polyhedron [Gale, Klee and Rockafellar, 1968]. Under the conditions we have imposed on the problem (2.2), we can prove a much stronger continuity condition which among other things allows us to show in the next section that the deterministic equivalent program (3.2) possesses strong regularity properties.

Theorem (7.7). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Suppose $z(x)$ is bounded on K. Then $z(x)$ satisfies a *Lipschitz* condition, i.e., there is some constant $\bar{B}$ such that x, x^0 in K imply

$$|z(x) - z(x^0)| \leq \bar{B} \, \| x - x^0 \|$$

where $\|\cdot\|$ denotes the Euclidean norm in R^n.

Proof. Again the linearity of $\bar{c}x$ allows us to restrict our attention to $Q(x)$. We must show that if $Q(x) > -\infty$ on K_2 and $x, x^0 \in K_2$ then there exists some constant B such that

$$|Q(x) - Q(x^0)| \leq B \, \| x - x^0 \| .$$

Given any x in K_2 and ξ in $\tilde{\Xi}$, $Q(x,\xi)$ is finite by Theorem (7.6) and can be expressed in terms of a basic solution of a linear program by writing

$$Q(x,\xi) = q(\xi)_{(i)} W^{-1}_{(i)} (p(\xi) - T(\xi)x) \tag{7.8}$$

where $W_{(i)}$ is a square nonsingular submatrix of W of rank m and $q(\xi)_{(i)}$ is the corresponding subvector of $q(\xi)$ as in theorem (4.1). Since by proposition (7.3) the function $Q(x,\xi)$ is convex polyhedral in x, the function

$$\phi(x, x^0; \xi) = \frac{|Q(x,\xi) - Q(x^0,\xi)|}{\| x - x^0 \|}$$

achieves its maximum when x and x^0 both belong to the region of linearity which has maximum slope. For x and x^0 in this region by (7.8) we have that there exist some index (i) such that

$$(7.9)\qquad |Q(x,\xi) - Q(x^0,\xi)| = |q(\xi)_{(i)}W^{-1}_{(i)}T(\xi)(x^0-x)|$$
$$\le B_{(i)}\,\|q(\xi)_{(i)}\|\cdot\|T(\xi)\|\cdot\|x^0 - x\|$$

where $B_{(i)}$ is some constant related to the determinant of $W^{-1}_{(i)}$. By the covariance condition (2.1) the right-hand side of the inequality is integrable. Since the integral is order preserving we have that

$$|Q(x) - Q(x^0)| \le \int |Q(x,\xi) - Q(x^0,\xi)|\,dF(\xi)$$
$$\le B_{(i)}\,\|x^0 - x\|\int \|q(\xi)_{(i)}\|\cdot\|T(\xi)\|\,dF(\xi) \le B\|x^0 - x\|$$

where B is the maximum over (i) of $B_{(i)}\int\|q(\xi)_{(i)}\|\cdot\|T(\xi)\|dF(\xi)$. B is finite since by theorem (7.2) there are only a finite number of cells in $\mathcal{K}$ and for each index (i) the finiteness of the integral is assured by the covariance condition (2.1).

So far we have made very little use of the form of the distribution of ξ to derive the properties of the objective of the deterministic equivalent program (3.2), except naturally for the covariance condition. Below, we obtain two interesting characterizations of $Q(x)$, or equivalently of $z(x)$, which rely on some further properties of $F(\xi)$. The first one of these theorems (7.17) can be obtained in many ways; for example, one could make use of the representation of the epigraph of $Q(x,\xi)$ in terms of its supporting hyperplanes, i.e., use the properties of the polar matrix of $\begin{bmatrix} q & 1 \\ W & 0 \end{bmatrix}$ [Wets, 1967], or one could also use some arguments involving the generalized inverse of W. However, in order to unify our treatment, we shall rely on Theorem (7.15), which is a mild generalization of a lemma in [Dantzig and Madansky, 1961].

A pair $(x,\xi) \in K_2 x\tilde{\Xi}$ is said to be *acceptable* if $Q(x,\xi)$ is finite. In view of Theorem (7.7) it is pointless to consider any other situation, and thus, for the remainder of this section we shall limit ourselves to this case. From duality theory, it follows that the linear program

$$\text{(7.10)} \qquad \begin{array}{ll} \text{Maximize} & w = \pi[p(\xi) - T(\xi)x] \\ \text{subject to} & \pi W \leq q(\xi) \end{array}$$

where π is an $\bar{m}$-row vector, is feasible and bounded whenever the pair (x,ξ) is acceptable. Let

$$D(\xi) = \{\pi | \pi W \leq q(\xi)\}$$

be the polyhedron determined by the constraints of (7.10). Its set of vertices will be denoted by $\text{ext } D(\xi)$. Note that $\text{ext } D(\xi)$ might be empty, but if it is empty for some ξ, then it is also empty for all ξ in $\tilde{\Xi}$. If this is the case and since the programs (7.10) are bounded for $(x,\xi) \in K_2 x\tilde{\Xi}$ we must then be confronted with an acute infection of degeneracy.

For each acceptable pair, let $\Pi(x,\xi)$ be the set of optimal solutions to (7.10), i.e., for each $\pi(x,\xi)$ in $\Pi(x,\xi)$ we have that $\pi(x,\xi)[p(\xi) - T(\xi)x] = Q(x,\xi)$. The set $\Pi(x,\xi)$ is the convex hull of vertices and extreme rays of $D(\xi)$. By $\Pi^{\dagger}(x,\xi)$ we denote the convex hull of the vertices of $D(\xi)$ contained in $\Pi(x,\xi)$. Note that $\Pi^{\dagger}(x,\xi)$ is empty only if $\text{ext } D(\xi)$ is empty.

Proposition (7.11). For each x in K_2, there exists a countable family $\{\pi(x,\xi)\}$ with domain $\tilde{\Xi}$ and range $R^{\bar{m}}$, with $\pi(x,\xi)$ $\mathcal{G}$-measurable and such that $\{\pi(x,\xi)\}$ is dense in $\Pi(x,\xi)$ for all ξ in $\tilde{\Xi}$. Such a function $\pi(x,\xi)$ will be called an $\mathcal{G}$-measurable selector.

Proof. For each x in K_2, the graph $\{(\Pi(x,\xi),\xi)\,|\,\xi \in \tilde{\Xi}\}$ of the multivalued function $\Pi(x,\xi)$ is a Borel subset of $R^{\bar{m}} \times \tilde{\Xi}$. This follows immediately from

(i) $\Pi(x,\xi)$ can be described by a finite number of algebraic expressions in $p(\xi), T(\xi)$ and $q(\xi)$ which are obtained from the optimality and feasibility requirements for linear programs;

(ii) $\tilde{\Xi}$ is a closed subset of R^N; and

(iii) $p(\xi), T(\xi)$ and $q(\xi)$ are coordinates (projections) of ξ.

Now, note that for all ξ in $\tilde{\Xi}$, the sets $\Pi(x,\xi)$ are non-empty and closed. Since $\mathcal{G}$ is the completion of the Borel algebra on R^N, a theorem on measurable selections [Castaing, 1967, Theorem (5.4)] yields the existence of the $\mathcal{G}$-measurable selectors $\{\pi(x,\xi)\}$.

For the remainder of this section, all we need is the existence of *some* $\mathcal{G}$-measurable selector $\pi(x,\xi)$ in $\Pi^{\dagger}(x,\xi)$. This can be obtained by invoking some weaker result on measurable selectors [Freedman, 1966; Theorem (4)] or by relying on the fact that in this case each function $\pi(x,\xi)$ is the convex combination of a finite number of "extreme" functions $\pi(x,\xi)$ passing through the vertices of polytopes determined by $\Pi^{\dagger}(x,\xi)$. A constructive but rather lengthy proof of the existence of some $\mathcal{G}$-measurable selector, can also be found in [Kall, 1967, Section 1, Satz 1].

Proposition (7.12). Consider an acceptable pair $(\bar{x},\xi)$. Suppose that $D(\xi) \neq 0$ and let $\pi(x,\xi) \in \Pi^{\dagger}(x,\xi)$. Then the hyperplane $H = \{(z,x) \mid z + \pi(\bar{x},\xi)T(\xi)x = \pi(\bar{x},\xi)p(\xi)\}$ is a supporting hyperplane of the epigraph $\{(z,x) \mid z \geq Q(x,\xi),\ x \in K_2\}$ of the convex polyhedral function $Q(x,\xi)$ at the point $(Q(\bar{x},\xi),\bar{x})$. Moreover, $\|\pi(\bar{x},\xi)T(\xi)\| \leq B(\xi)$ where $B(\xi)$ is the maximum slope of a linear part of the function $Q(x,\xi)$.

Proof. The first part of the proposition follows from the fact that for all x in K_2 and ξ in $\tilde{\Xi}$, $\pi(\bar{x},\xi)$ is a feasible but not necessarily optimal solution of corresponding linear program (7.10), and thus

$$Q(x,\xi) = \pi(x,\xi)[p(\xi) - T(\xi)x] \geq \pi(\bar{x},\xi)[p(\xi) - T(\xi)x].$$

Hence, for all x in K_2 and $z \geq Q(x,\xi)$, i.e., for the points belonging to the epigraph of $Q(x,\xi)$, we have that

$$z + \pi(\bar{x},\xi)T(\xi)x \geq \pi(\bar{x},\xi)p(\xi). \tag{7.13}$$

That H supports at $(Q(\bar{x},\xi),\bar{x})$, follows from the identity

$$Q(\bar{x},\xi) = \pi(\bar{x},\xi)[p(\xi) - T(\xi)\bar{x}]. \tag{7.14}$$

The remainder follows from the observation that $\pi(x,\xi)$ is a vertex or a convex combination of vertices of $D(\xi)$, which correspond to basic solutions of (7.10). Arguments similar to those invoked to obtain (7.9) yield the desired inequality.

The first part of the following theorem is a slight generalization of Lemma 2 in [Dantzig and Madansky, 1961]. Although it is a trivial consequence of the two previous propositions, it has proven to be extremely useful as a conceptual guide when one seeks algorithmic procedures for solving stochastic programs.

Theorem (7.15). Let $\bar{x}$ belong to K_2 and $\pi(\bar{x},\xi)$ is an $\mathcal{G}$-measurable selection of $\Pi^{\dagger}(\bar{x},\xi)$. Then the hyperplane $\{(z,x)\,|\,z + E_{\xi}\{\pi(\bar{x},\xi)T(\xi)\}x = E_{\xi}\{\pi(\bar{x},\xi)p(\xi)\}\}$ is a supporting hyperplane of the epigraph $\{(z,x)\,|\,z \geq Q(x), x \in K_2\}$ of $Q(x)$ at the point $(Q(\bar{x}),\bar{x})$. Moreover, $E_{\xi}\{||\pi(\bar{x},\xi)T(\xi)||\} \leq B$ where B is the Lipschitz constant of $Q(x)$ as in (7.7).

Proof. The first part of the theorem follows from integrating both sides of the relations (7.13) and (7.14). The measurability question is taken care of by the fact that in each case the expressions involved are continuous functions of measurable function and the finiteness of the integrals is assured by our assumptions on problem (7.10) and the covariance condition (2.1). As for the remainder, it follows straightforward from (7.9) and the definition of $\Pi^{\dagger}(\bar{x},\xi)$.

Corollary (7.16). Suppose the matrix T is fixed. For all $\bar{x}$ in K_2, let $\pi(\bar{x}) = E_{\xi}\{\pi(\bar{x},\xi)\}$ where $\pi(\bar{x},\xi)$ is an $\mathcal{G}$-measurable selection. Then the vector $(1,\pi(\bar{x})T)$ is the normal of a supporting hyperplane of the epigraph of $Q(x)$ at $(Q(\bar{x}),\bar{x})$. Moreover, $||\pi(\bar{x})T|| \leq B$.

Theorem (7.17). If $\tilde{\Xi}$ is a finite set, i.e., when $F(\xi)$ is a finite discrete distribution, then $Q(x)$ is a convex polyhedral function.

Proof. $Q(x) = E_{\xi}\{Q(x,\xi)\}$ is by (7.5) a convex combination of a finite number of polyhedral functions in x.

Combining the last theorem with the second part of theorem (4.7) it follows naturally that when $\tilde{\Xi}$ is finite the deterministic equivalent program (3.2) can be written as the minimization of a convex polyhedral function on a convex polyhedron or by introducing some additional constraints as a linear program. One way to do so, but not necessarily the most efficient one (especially if the assumption of relatively complete recourse is not satisfied and some of the components of q and T are random), is to express the problem as a large-scale linear program along the lines of [Dantzig and Madansky, 1961; (29), (37)], a variant of which can be found in [Wets, 1966-b; Section 3B, Case 1].

If $F(\xi)$ is an absolutely continuous distribution, then the following propositions can be found in [Kall, 1967], and [Wets, 1966-b], respectively.

Proposition (7.18). Consider a stochastic program with complete recourse, such that q is fixed and $F(\xi)$ is an absolutely continuous distribution. Then $Q(x)$ is differentiable on $K_2 = R^n$.

Proposition (7.19). Consider a stochastic program with fixed recourse such that q and T are fixed and $F(\xi)$ is an absolutely continuous distribution; then $Q(x)$ is differentiable.

Both proofs rely essentially on the fact that for a given x, $\Pi(x,\xi)$ is only multivalued on sets of measure zero. Thus in view of theorem (7.15), integrating $\pi(x,\xi)T(\xi)$ on $\tilde{\Xi}$ where $\pi(x,\xi)$ is any $\mathcal{G}$-measurable selection always determines the same supporting hyperplane of the epigraph of $Q(x)$. From this it is implied that the supports are unique for all x in K_2 i.e., the convex function $Q(x)$ is differentiable [Rockafellar, 1969; Section 25]. Although both propositions are correct, their proofs are incomplete since they fail to show that every normal to a support of $Q(x)$ can be obtained as the integral of $\mathcal{G}$-measurable selectors of $\Pi(x,\xi)$. This and some generalizations of the above propositions will be included in a projected paper.

For stochastic programs with simple recourse, it suffices that the marginal distribution of the subvector of random variables $(p_i(\xi), T_i(\xi), q_i(\xi))$ $i = 1,\ldots,\bar{m}$ be absolutely continuous [Walkup and Wets, 1969-a; Proposition (2.8)]. If only the right-hand sides $p(\xi)$ are random, it suffices that each marginal distribution of ξ be continuous [Wets, 1966-a; Proposition (21)]. All these constitute sufficient conditions for differentiability of $Q(x)$. There does not seem to be any simple condition which would also be necessary for the differentiability of $Q(x)$.

8. <u>Some Regularity Properties of the Equivalent Program</u>. The study of constrained optimization has led to various regularity conditions for optimization problems which in some sense determine if the problem is "well" formulated and usually give some indication as to the type of method(s) one could reasonably expect to generate solution procedures. The standard

approach is to study the effect that small perturbations of some of the constraints will have on the optimum. For *convex programs* these regularity conditions have been traditionally--and for good reasons--related to the properties of a so-called dual problem, although they can very often be verified without necessarily deriving the actual dual problem [Rockafellar, 1967], [Van Slyke and Wets, 1968], [Rockafellar, 1968].

Definition (8.1). A convex program is

(i) *feasible* if the set determined by the constraints is nonempty,

(ii) *solvable* if the value of the infimum is finite and achieved for some value of the variable,

(iii) *dualizable* if there is no duality gap, i.e., if the optimal value of the convex program and its dual are equal,

(iv) *stable* if there exist (optimal) nontrivial Lagrange multipliers or equivalently for convex programs if the dual problem is solvable.

The terminology used here differs from that used by [Rockafellar, 1967] and [Van Slyke and Wets, 1968] in only one respect: The dualizable condition corresponds to what they refer to as normality. Note that if a convex program is stable, it is also dualizable [Van Slyke and Wets, 1968; Proposition (6.8)] and obviously that solvability implies feasibility.

In Sections 4 and 7 we have established that the deterministic equivalent program of a stochastic program with fixed recourse whose random elements satisfy a covariance condition is a convex program of the form

(8.2) $$\begin{aligned} \text{Find } \inf z(x) &= \bar{c}x + Q(x) \\ Ax &= b \\ W^*T(\zeta)x &\geq W^*p(\zeta) \quad \forall \zeta \in \tilde{\Xi}_{p,T} \\ x &\geq 0 \end{aligned}$$

where $z(x)$ is a convex function which is either identically $-\infty$ on the subset K_2 of R^n determined by the induced constraints or it is finite on K_2 in which case it is Lipschitzian with constant $\bar{B} = ||c|| + B$, where B is the constant for Q as defined by Theorem (7.7). Here, we shall consider perturbations of the fixed constraints $Ax = b$. Following [Van Slyke and Wets, 1968; Section 3] we can thus write the dual program of (8.2) as

(8.3) $$\begin{aligned} &\text{Find } \sup \nu \\ &\text{subject to } \nu \leq Q(x) + (c-\pi A)x + \pi b \\ &\qquad\qquad \text{for all } x \in K_2 \quad \{x \mid x \geq 0\}. \end{aligned}$$

This program can be interpreted as seeking the "highest" supporting hyperplane of the epigraph $\mathcal{C}$ of the variational function $\phi(u) = \{\operatorname{Inf} z(x) \mid Ax = b - u,\ x \in K_2 \cap R^n_{\oplus}\}$.

Characterizing feasibility for the program (8.2) has been the burden of Section 5. In the remainder of this section we investigate some sufficient conditions for the program (8.2) to satisfy one or more of the properties listed in definition (8.1). We shall see that for a *broad* class of stochastic programs with fixed recourse, the equivalent convex program possesses the desirable regularity properties.

Theorem (8.4). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition and suppose that the constraint set K is bounded, then the deterministic equivalent program (8.2) is solvable and dualizable.

Proof. If $z(x) = -\infty$ on K, the theorem follows trivially from the standard $\pm\infty$ conventions. If $z(x)$ is finite, it is Lipschitz (7.7) and thus continuous, i.e., attains its minimum on $K = K_1 \cap K_2$ which is compact since it is bounded by assumption and closed since it is the intersection of a closed polyhedron K_1 and the closed set K_2 (Theorem (4.7)). Thus (8.1) is solvable. It remains to be shown that in this case it is also dualizable. This follows in a rather straightforward fashion from [Van Slyke and Wets, 1968; Proposition (5.1)] and the observation that the compactness of K and the continuity of $z(x)$ is sufficient to establish that the epigraph $\mathcal{C}$ of the variational function $\phi(y) = \{\text{Min } z(x) \mid x \in K-y\}$ is a closed set.

Easy examples of stochastic programs with recourse can be found, satisfying all hypotheses of the previous theorem except the boundedness of K, whose infimum is finite but which are not solvable. In [Williams, 1965] one can find a detailed characterization of this situation for stochastic program with simple recourse. The dualizability of (8.2) also requires some restrictions. The following example which has only T random and K unbounded possesses a deterministic equivalent program which is not dualizable; in fact, in this case there is an infinite duality gap.

<u>Example (8.5)</u>. The deterministic equivalent program of

$$\begin{array}{lllll} \text{Find} \quad \inf z(x) = & -x_2 + E_\xi\{\text{Min } y_1 & & & \} \\ & x_1 - x_2 & & & = 0 \\ & \xi_1 x_1 & - y_1 - y_2 & & = 0 \\ & \xi_2 x_2 & & - y_2 - y_3 & = 0 \\ & x_1,\ x_2 \geq 0 & y_1,\ y_2,\ y_3 \geq 0 & & \end{array}$$

where ξ_1 has a continuous distribution on $[1,\infty)$ with density $f(\xi_1) = \xi_1^{-2}$ and $\xi_2 = \xi_1 - 1$, has an optimal value of 0. The optimum value of the dual of the deterministic equivalent program is $-\infty$. It is not known if there exist stochastic programs (with fixed recourse or not) whose deterministic equivalent programs exhibit a finite duality gap.

The following lemma proved in [Walkup and Wets, 1969-d] is particularly useful since it is immediately applicable to all stochastic programs whose constraints determine a polyhedral region. Theorems (4.7), (4.10) and Corollary (4.13) have shown that this will be the case in all but the most sophisticated applications.

Lemma (8.6). Consider the mathematical program

$$(8.7) \qquad \begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && Ax = b, \\ &&& x \geq 0 \end{aligned}$$

where the objective f is convex and Lipschitz on a polyhedron and (8.7) is finite, then (8.7) is stable.

Theorem (8.8). Consider a stochastic program with fixed recourse whose random elements satisfy a weak covariance condition. Suppose that K is polyhedral and the program is finite. Then the convex program (8.2) is stable.

Proof. It suffices to observe that by Theorems (7.6) and (7.7) the hypotheses of the Lemma (8.6) are satisfied.

To see that the hypotheses of Theorem (8.8) are necessary to obtain stability, consider again the example (8.5) with the additional fixed constraint $x_1 = 1$. This problem has a finite optimum, in fact, is solvable. However, small perturbations of the constraints will change the value of the optimum drastically. It is easy to verify that the dual is not solvable.

Acknowledgements

The manuscript has been read by Mark Eisner, Robert Fricks, and David Walkup, whose constructive criticisms have been of great value in preparing this text.

References

M. Avril and A. C. Williams, 1969: "The Value of Information and Stochastic Programming," Manuscript, Socony Mobil Oil (submitted to *Oper. Res.*).

E. Beale, 1955: "On Minimizing a Convex Function Subject to Linear Inequalities," *Journal of the Royal Statistical Society*, Series B 17, 173-184.

W. Bonnice and V. Klee, 1963: "The Generation of Convex Hulls," *Math. Ann.* 1952, 1-29.

Ch. Castaing, 1967: "Sur les multi-applications mesurables." Thèse présentée à la Faculté des Sciences de l'Université de Caen.

A. Charnes, W. Cooper and G. Thompson, 1965: "Constrained Generalized Medians and Hypermedians as Deterministic Equivalents for Two-Stage Linear Programs under Uncertainty," *Management Sci.* 12, 83-112.

Choquet, 1956: "Unicité des représentations intégrales au moyens des points extrémaux dans les cônes convexes réticulés," *C. R. Acad. Sci. Paris* 243, 555-557; "Existence des représentations intégrales au moyen des points extrémaux dans les cônes convexes," *C. R. Acad. Sci. Paris* 243, 699-702.

G. Dantzig, 1955: "Linear Programming under Uncertainty," *Management Sci.* 1, 197-206.

G. Dantzig, 1963: "Linear Programming and Extensions, Princeton University Press, Princeton.

G. Dantzig and A. Madansky, 1961: "On the Solution of Two-Stage Linear Programs under Uncertainty," in *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 165-176.

C. Davis, 1954: "Theory of Positive Linear Independence," *Amer. J. Math.* 76, 733-746.

M. Dempster, 1968: "On Stochastic Programming: I. Static Linear Programming under Risk," *J. Math. Anal. Appl.* 21, 304-343.

H. Eggleston, B. Grunbaum and V. Klee, 1964: "Some Semi-Continuity Theorems for Convex Polytopes and Cell-Complexes," *Commentarii Mathematici Helvetici* 39, 165-188.

M. El-Agizy, 1967: "Two-Stage Programming under Uncertainty with Discrete Distribution Function," *Oper. Res.* 15, 55-70.

A. Ferguson and G. Dantzig, 1956: "The Allocation of Aircraft to Routes: An Example of Linear Programming under Uncertain Demand," *Management Sci.* 3, 45-73.

D. Freedman, 1966: "On the Equivalence Relations between Measures," *Annals of Math. Statist.* 37, 686-689.

R. Fulkerson, 1968: "Blocking Polyhedra," Memorandum RM-5834-PR, RAND Corporation, Santa Monica, California.

D. Gale, V. Klee and R. Rockafellar, 1968: "Convex Functions and Convex Polytopes," *Proc. Amer. Math. Soc.* 19, 867-873.

B. Grünbaum, 1967: "Convex Polytopes, Wiley, New York.

W. Hansen and V. Klee, 1969: "Intersection Theorems for Positive Sets," *Proc. Amer. Math. Soc.* 22, No. 2, 450-457.

P. Kall, 1967: "Das zweistufige Problem der stochastichen linearen Programmierung," *Z. Wahrscheinlichkeitstheorie und verw. Geb.* 8, 101-112.

V. Klee, 1957: "Extremal Structures of Convex Sets," *Arch. Math.* 8, 234-240.

A. Madansky, 1960: "Inequalities for Stochastic Linear Programming Problems," *Management Sci.* 6, 197-204.

R. McKinney, 1962: "Positive Bases for Linear Spaces," *Trans. Amer. Math. Soc.* 103, 131-148.

K. Miyasawa, 1968: "Information Structures in Stochastic Programming Problems," *Management Sci.* 14, 275-291.

S. Parrikh, 1967: "Generalized Stochastic Programs with Deterministic Recourse," Operations Research Center ORC 67-27, University of California, Berkeley.

J. Reay, 1965-a: "Generalization of a Theorem of Caratheodory, *Amer. Math. Soc. Memoir*, No. 54.

J. Reay, 1965-b: "A New Proof of the Bonnice-Klee Theorem," *Proc. Amer. Math. Soc.* 16, 585-587.

R. T. Rockafellar, 1967: "Duality and Stability in Extremum Problems Involving Convex Functions," *Pacific J. Math.* 21, 167-187.

R. T. Rockafellar, 1968: "Duality in Nonlinear Programming," in *Mathematics of Decision Sciences, Lectures in Applied Mathematics*, Vol. 11, Providence, R.I.

R. T. Rockafellar, 1969: "Convex Analysis," Princeton University Press, Princeton, N.J.

J. Sengupta, G. Tintner and C. Millham, 1963: "On Some Theorems in Stochastic Linear Programming with Applications," *Management Sci.* 10, 143-159.

G. Tintner, 1960: "A Note on Stochastic Linear Programming," *Econometrica* 28, 490-495.

R. Van Slyke and R. Wets, 1968: "A Duality Theory for Abstract Mathematical Programs with Applications to Optimal Control Theory," *J. Math. Anal. Appl.* 22, 679-706.

R. Van Slyke and R. Wets, 1969: "L-Shaped Linear Programs with Applications to Optimal Control and Stochastic Linear Programming," *SIAM J. on Appl. Math.* 17.

D. Walkup and R. Wets. 1967-a: "Continuity of Some Convex-Cone Valued Mappings," *Proc. Amer. Math. Soc.* 18, 229-235.

D. Walkup and R. Wets, 1967-b: "Stochastic Programs with Recourse," *SIAM J. on Appl. Math.* 15, 1299-1314.

D. Walkup and R. Wets, 1968: "A Note on Decision Rules for Stochastic Programs," *J. Computer and System Sci.* 2, 305-311.

D. Walkup and R. Wets, 1969-a: "Stochastic Programs with Recourse: Special Forms," *Proceedings of Sixth International Symposium on Mathematical Programming*, H. W. Kuhn, Ed., to appear.

D. Walkup and R. Wets, 1969-b: "Lifting Projections of Convex Polyhedra," *Pacific J. Math.* 28, 465-475.

D. Walkup and R. Wets, 1969-c: "Stochastic Programs with Recourse II: On the Continuity of the Objective," *SIAM J. Appl. Math.* 17, 98-103.

D. Walkup and R. Wets, 1969-d: "Some Practical Regularity Conditions for Nonlinear Programs," *SIAM J. on Control* 7.

R. Wets, 1966-a: "Programming under Uncertainty: The Complete Problem," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 4, 316-339.

R. Wets, 1966-b: "Programming under Uncertainty: The Equivalent Convex Program," *SIAM J. Appl. Math.* 14, 89-105.

R. Wets, 1966-c: "Programming under Uncertainty: The Solution Set," *SIAM J. Appl. Math.* 14, 1143-1151.

R. Wets, 1967: "Notes on Stochastic Programming," Lecture Notes, University of California, Berkeley (unpublished).

R. Wets, 1968: "On a Paper by Charnes, Kirby, and Raike," Boeing Scientific Research Laboratories Mathematical Note No. 550.

R. Wets and C. Witzgall, 1967: "Algorithms for Frames and Lineality Spaces of Cones," *J. Res. NBS, Math. and Math. Phys.* 71B, 1-7.

R. Wets and C. Witzgall, 1968: "Towards an Algebraic Characterization of Convex Polyhedral Cones," *Numerische Mathematik* 12, 134-138.

A. Williams, 1963: "A Stochastic Transportation Problem," *Oper. Res.* 11, 759-770.

A. Williams, 1965: "On Stochastic Linear Programming," *SIAM J. Appl. Math.* 13, 927-940.